Novel Multiple Lane Detection Based on Dense Vanishing Point Estimation

Umar Ozgunalp, Xiao Ai, and Naim Dahnoun

***Abstract*—The detection of multiple curved lanes is still a challenge for Driver Assistance Systems (DAS) today, due to interferences such as road markings and shadows casted by road side structures and vehicles. The vanishing point (V_p) contains the global information of the road image. Hence, V_p based lane detection algorithms are quite insensitive to interference. However, when curved lanes are assumed, V_p shifts with respect to the rows of the image. In this paper, a V_p for each individual row of the image is estimated by first extracting a V_{py} (vertical position of the V_p) for each individual row of the image from v-disparity. Then, based on estimated V_{py}s a 2D V_{px} accumulator is efficiently formed. Thus, by globally optimizing this 2D V_{px} accumulator, globally optimum V_ps for road image is extracted. Then, estimated V_ps are utilized for multiple curved lane detection on non-flat surfaces. The resultant system achieves a detection rate of 99% in 1277 frames of 5 stereo vision test sequence.**

***Index Terms*—Lane detection, Stereo vision, v-disparity, Dynamic programing, Vanishing point detection.**

I. Introduction

ACCORDING to statistics [1], around 70% of all reported road accidents in Great Britain are a result of driver error or slow reaction time. Fortunately, the computation power available today makes it possible to utilize DAS to prevent or minimize the consequences of these accidents. By using specialized algorithms, DAS predicts driver intent, warns the driver about possible lane departure or collision, as well as many more functionalities.

Lane detection is one of the key elements of DAS [2] and it is necessary for lane departure warning. Due to the changing environment, the input image can be noisy and lane detection can be a challenging task. For example, changing light conditions or the lack of consistent painting can affect the lane detection significantly. Thus, some assumptions are commonly made in the algorithms to increase the performance such as constant road width, constant lane painting width, consistent road texture and a flat road [3]. An important property of the input image is perspective mapping. During the image capturing process, the vision sensor maps the three dimensional world information into a two dimensional image. During this process, all parallel lines in the world coordinate system converge on a point (V_p) in the image coordinate system. Under the assumption that the lanes are parallel to each other, V_p can be used to improve the system robustness significantly. Thus, many researchers [4]–[8] are focused on V_p based lane detection algorithms since, V_p contains global information and V_p based algorithms are less sensitive to local noise such as, occlusion or shadows. Previously developed V_p based lane detection algorithms demonstrated robust results by first detecting V_p and then detecting lanes based on this global information. However, they still have limitations, such as having flat road assumption, a straight lane model and the ability of to detect only the current lane.

In [6], the algorithm assumes intrinsic and extrinsic parameters of the camera are known and that the vehicle is travelling parallel to the road. Thus, V_p can be estimated from these parameters and lines crossing V_p can be searched by using the $2D$ Hough transform. In [4] and [5], the algorithms first detect and track V_p (including the horizon line). Therefore, this eliminates the assumptions of knowing extrinsic camera parameters and the assumption of the vehicle travelling parallel to the road. As a second step, algorithms search for line pairs crossing V_p for each lane (a lane is a light stripe on a darker background. Thus, each lane has two boundaries).

With a single V_p, only linear lane models can be used and algorithms using a single V_p are only suitable for the roads with limited curvature, such as motorways. In [7], the algorithm segments the image into horizontal bands and detects a V_p (also assuming the horizon line is known) for each band. The algorithm first detects a V_p in a bottom band and then moves to the upper band. Depending on the previously detected V_p, the algorithm sets a search range for the V_p which appears in the upper band and detects a new V_p for this band. Thus, the algorithm detects a few V_ps iteratively. There are two problems with this approach. The first issue is that the algorithm treats the lanes as they are piecewise linear within a few bands. Thus, the accuracy of V_p decreases in the areas close to the band boundaries. The more important issue with this approach is it is no longer global for the image. To detect V_p in the bottom band (or in the other bands), only a small portion of the image is used and the rest is ignored. For example, even though there is enough information in the complete image to detect V_p accurately, if the information in the bottom band is too noisy, the algorithm can fail. In [8], the CHEVP algorithm (Canny/Hough Estimation of Vanishing Points) is proposed to initialize lanes. CHEVP also segments the image into a few bands and iteratively detects a V_p for each band (similarly to [7]). By detecting V_ps, the direction and curvature of the parallel lanes can be estimated. However, information such as lateral offset still needs to be detected. This algorithm initializes both the left and the right lanes by using the straight lines extracted from the Hough transform.

U. Ozgunalp, Xiao Ai, and N. Dahnoun are with Department of Electrical and Electronic Engineering, University of Bristol Merchant Venturers Building, Woodland Road, BS8 1UB, Bristol, U.K e-mail: (Umar.Ozgunalp@bristol.ac.uk , Xiao.Ai@bristol.ac.uk, Naim.Dahnoun@bristol.ac.uk)).

Manuscript received ; revised .

It selects two lines, directed to the V_p, from the bottom band (if available) which are most close to the center column of the image. This approach of selecting lines is rather simplistic and highly dependent on thresholds. Any road mark on the near field of the road, directed to the V_p, may also cause the algorithm to fail.

Existing V_p based lane detection algorithms either use single [4]–[6] or few [7], [8] V_{px} (the horizontal position of the Vp) on a single horizon line. While single V_p based algorithms are based on only the linear lane model, multiple V_p based algorithms utilize a non-global iterative approach to detect multiple V_ps. Existing algorithms also estimate a single horizon line (a planar road assumption) using a single image, which is prone to noise such as camera shakes. However, using stereovision to estimate the horizon line can be a more robust method, such as proposed in [9].

In this paper, a global way to estimate V_p (both in horizontal direction (V_{px}) and in vertical direction (V_{py})) for each individual row of the image is proposed. Then, this V_p curve is utilized for multiple curved lane detection on non-flat surfaces. A block diagram of the proposed system is illustrated in Fig. 1. In this paper, we used stereo vision as the input. Stereovision reveals valuable 3D world information to extract the vertical road profile for non-flat roads. Using the disparity map, first the vertical profile of the road is extracted. Then, since the vertical profile of the road is already known, any feature point which does not appear on the road is eliminated. Secondly, a horizon line for each individual row of the road image is calculated using the estimated vertical road profile. At this stage, since the horizon line is already known, a V_{px} is calculated for each individual row of the image using a global optimization technique in an efficient way. Thirdly, based on the known V_ps, a 1D likelihood accumulator is created and peak pairs (due to the dark-light-dark transition of lanes) are searched in this accumulator to detect the lanes. As a final step of the algorithm, the change in the lateral offset is estimated by cross correlating estimated 1D likelihood functions for consecutive frames. Then, based on the estimated lateral offset change, 1D likelihood signals are combined for consecutive frames to improve the signal to noise ratio (SNR).

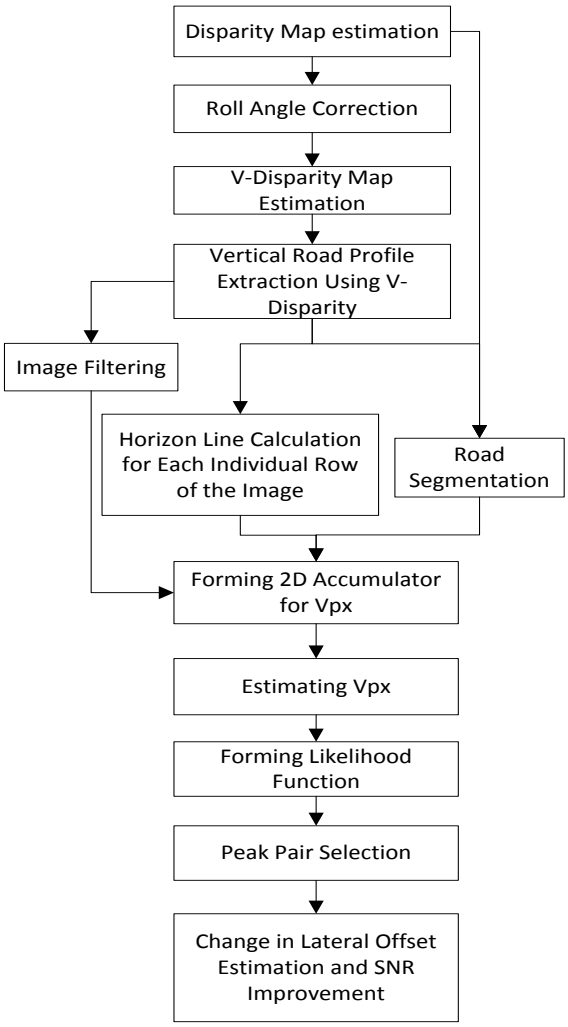
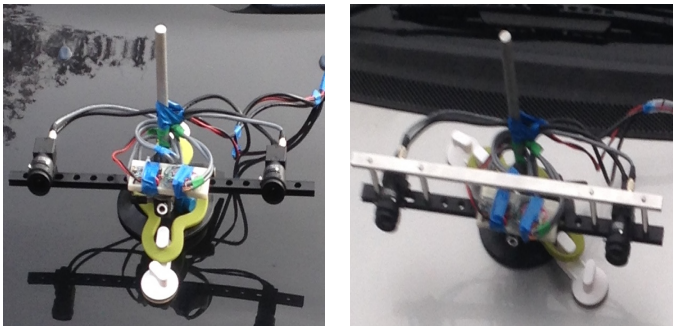


Fig. 1: Block diagram of the system

II. Road Profile Extraction

A. Experimental Set-Up

In our stereo camera rig, two Point Grey Flea3 (FL3-GE-13S2C-CS) cameras have been used. These cameras have 3.75 μm sensors and they can capture up to 1.3 MP images with 31 fps. Synchronization has been achieved by triggering the cameras using a pulse width modulation signal (the same signal for both of the cameras) using an Arduino board. The base line of the cameras is set to 34 cm. Example set-ups are illustrated in Fig. 2.



(a) (b)

Fig. 2: Example experimental set ups

B. Disparity Map Estimation

The initial component of the algorithm is the disparity map estimation [10]. The disparity map estimation, which outputs the 3D world information, is useful for both extracting the vertical profile of the road and segmenting it. Although, there

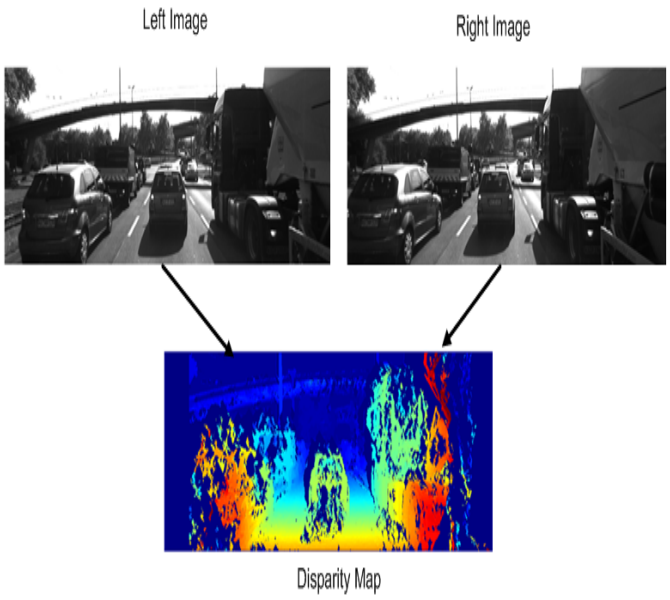


Fig. 3: Disparity map estimation from stereo images

are several applicable stereo vision algorithms available in the literature, only a limited number of algorithms can achieve good accuracy while working in real-time, such as [11] and [12]. In this paper, we have used our previously published algorithm [13] to acquire the disparity map. This algorithm is suitable for our application due to its good accuracy and high computational efficiency. In Fig. 3, input stereo images and their corresponding calculated disparity maps are illustrated.

C. Roll Angle correction and v-disparity Map

In the previous section, 3D world information is extracted by stereo vision. Then, this information can be used for extracting the road profile via the use of the v-dispairty map. The v-disparity map is a widely applied method for ground plane extraction [14], [15], [16]. The algorithm, creates a histogram of disparities for each row of the image and then maps them to the 2D v-disparity map.

Roll angle is generally assumed to be zero by many v-disparity based algorithms. However, due to the camera installation to the vehicle, it can vary significantly. Thus, a roll angle should be initially estimated for more robust results. In this paper, the roll angle is estimated by fitting a plane to a small patch from the near field in the disparity map. Then, using the estimated roll angle, both the input image and the disparity map are rotated using the affine transform. The roll angle is only estimated in the first frame and the same roll angle is used for the rest of the video sequence to decrease the computational complexity. With roll angle correction, it is seen that the algorithm can create a better v-disparity map and can work more robustly. In Fig. 4(a), original road image is shown. In Fig. 4(b), the disparity map of the image is shown and the patch used for the roll angle estimation is depicted by the black box. In Fig. 4(c), the rotated road image is shown. In Fig. 4(d), the rotated disparity is shown. In Fig. 4(e), the v-disparity map of the original image is shown and, in Fig. 4(f),

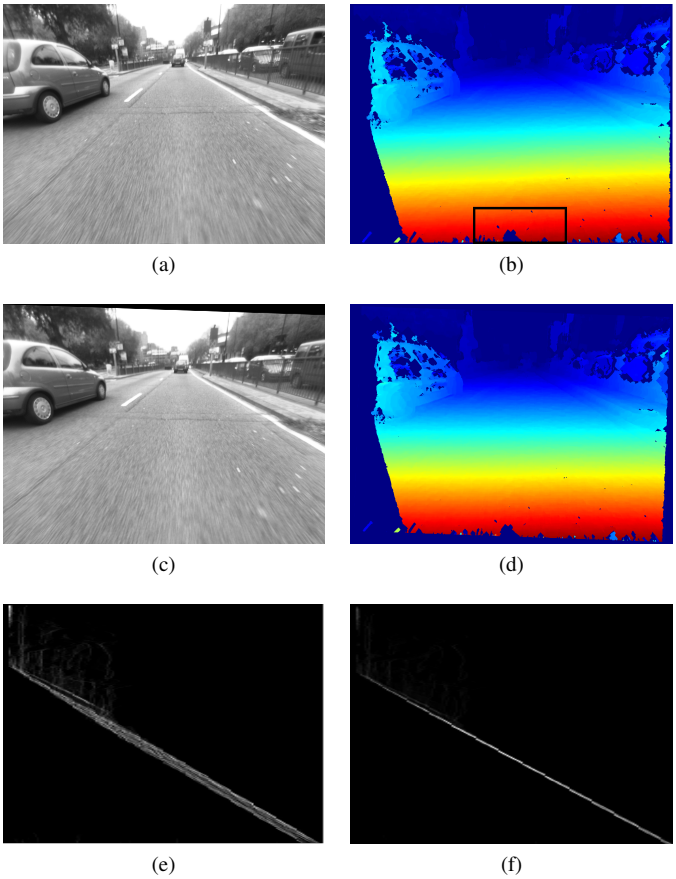


Fig. 4: Roll angle correction. (a) Input image, (b) Disparity map of the input image. The patch used for the roll angle estimation is depicted by the black box. (c) Road image after roll angle correction. (d) Disparity map after roll angle correction. (e) v-disparity map of the original image. (f) v-disparity map after roll angle correction.

the v-disparity of the rotated disparity map is shown. As is seen from Fig. 4(e) and Fig. 4(f), after the roll angle correction, the v-disparity map becomes more sharp. Thus, the vertical profile of the road can be estimated more accurately and the road can be segmented more precisely.

D. Energy Minimization Based on Dynamic Programming

For a flat road, the disparity of the road for each row of the image should decrease linearly. Thus, the road profile is projected to the v-disparity map as a straight line. As a robust line detector, the Hough transform [17] is a common method for extracting the road profile from the v-disparity map. However, with this technique, the vertical road profile can only be extracted under a flat road assumption or with limited flexibility. In our previous paper [15], we have successfully implemented such a technique and applied it for an obstacle detection application. In this paper, the accuracy requirements are high and this renders such an approach unsuitable. Thus, a more flexible approach has been taken. After creating the v-disparity map, dynamic programming [18] is adopted for extracting the road profile.

Fig. 5: Example road image with a lack of lane painting in the near field

Dynamic programming is a good approach to explore a path of minimized energy [19] under the interferences of a great number of outliers and a high level of noise.

In this paper, dynamic programming is utilized in a two-pass optimization. First, it optimizes (globally) the v-disparity map and estimates connected V_{py}s for each individual row of the image. The algorithm then uses the estimated V_{py} values to create a 2D V_{px} accumulator (see section II-H for details) in an optimized way. The proposed 2D V_{px} accumulator (V_{px} vs row number) is then optimized globally via dynamic programming for estimating connected V_{px}s for each row of the image. Therefore, a continuously varying V_p curve in both the vertical direction (V_{py}) and the horizontal direction (V_{px}) is optimized globally for curved lanes on non-flat surfaces.

Traditionally, for the V_{px} estimation of curved lanes, algorithms segment the image into a few horizontal image bands and iteratively, detect V_{px} of the bands from the bottom band to the upper bands (assuming the road is flat and, therefore, V_{py} is the same for all the bands). However, this approach is not global and, while estimating V_{px} of the current band, the algorithms completely ignore the information supplied in the upper bands. This non-global approach may lead to mis-detection. For example, mis-detection will occur, if there is a higher level of noise in the near field or lack of lane painting. An example of such a case is demonstrated in Fig. 5, where there is no lane painting in the near field. However, there is enough information in the complete image to detect the lanes.

In this paper, V_p is optimized globally by minimizing the energy function in equation 1. The data term, E_{data}, penalizes the disagreement in V_p. In other words, it depends on the total vote each accumulator gets. The road profile can be assumed to be piecewise smooth. Thus, V_p should also be piecewise smooth. The smoothness term, E_{smooth}, penalizes the change in V_p and ensures smoothness, where, λ is a constant.

$$E(r) = E_{data}(r) + \lambda E_{smooth}(r) \tag{1}$$

The first stage in which dynamic programming has been used is in optimizing the v-disparity map to extract the vertical profile of the road. Dynamic programming has been used to search a path in the v-disparity map which starts from the right and goes to the left (starting from the left would also give exactly the same result). Since the upper rows are further away from the camera, they should have either a decreasing disparity value or the same disparity value (the input disparity map does not have sub-pixel accuracy). Thus, only this pattern is searched.

Let d be an individual road disparity value for a row in the road image and r_d be the row number that disparity belongs to. Let $m(r_d)$ be the cost function (a column in the v-disparity map) for that disparity. Then, the E_{data} can be expressed as in equation 2.

$$E_{data}(r) = \sum_{d=dmax}^{1} m(r_d) \tag{2}$$

where $dmax$ is the maximum disparity value in the v-disparity map. Let $s(r_d, r_{d+1})$ be E_{smooth} and $s(r_d, r_{d+1}) = abs(r_d - r_{d+1})$, where r_{d+1} is the row number in the next disparity column (next to the one r_d belongs to) in the v-disparity. E_{smooth} penalizes the sharp changes between the assigned disparity values. Due to computational efficiency concerns, this is chosen to be the absolute difference. However, in some applications, it could be a more complex formula such as the spring equation. E_{smooth} can be expressed as in equation 3.

$$E_{smooth}(r) = \sum_{d=dmax-1}^{1} s(r_d, r_{d+1}) \tag{3}$$

where λ is a smoothness constant. By using equation 2 and equation 3, equation 1 can be rewritten as equation 4 and, then, as equation 5

$$E(r) = \sum_{d=dmax}^{1} m(r_d) + \lambda \sum_{d=dmax-1}^{1} s(r_d, r_{d+1}) \tag{4}$$

$$E(r) = m(r_{dmax}) + \sum_{d=dmax-1}^{1} m(r_d) + \lambda s(r_d, r_{d+1}) \tag{5}$$

Equation 5 can be solved iteratively. In each iteration, by minimizing equation 5 for a range of r, a relationship between r_d and r_{d+1} can be calculated and saved into a buffer with the same size of v-disparity map. Thus, once a global optimum is found for the final d, the algorithm can trace back to the beginning of the path. More details can be found in [18] and [19]. An example result is given in Fig. 6 and the estimated road profile is illustrated with the blue line on the top of the v-disparity map.

In the energy function, the smoothness term is also important for determining how far the road profile estimation can be considered as reliable. Many lane detection algorithms may give unreliable results for the far field or restrict the output results to the fixed arbitrary range, since the SNR drastically decreases when moving to the far field in the image. However, this range depends on many parameters such as the visible area of the road (this region may be occluded) and should be estimated by analysing the input signal. If there are not enough visible road points in the far field, there would not be a distinct line at the end of the v-disparity map. In this algorithm, if the far field is occluded, due to the smoothness term in equation 1, the estimated path would have a flattened region at the end. Thus, the unreliable area in the far field can be estimated.

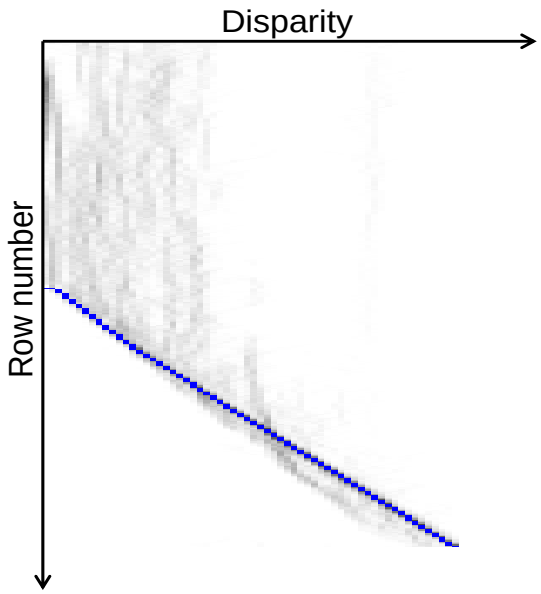


Fig. 6: V-Disparity map for vertical profile estimation

Once this region is removed from the estimated V_{py}s, the 1D output (a V_{py} for each row. i.e. the blue line in Fig. 6) is fitted to a quadratic. The advantages of the fitting stage, include the estimation of the road profile at the far field and reducing the output parameters from a few hundered (depends on the image resolution) to a few. Thus, tracking in the time domain would be much more efficient.

E. Speckle Noise Elimination

The system is tested under various conditions, including sunny and cloudy days. One would expect that the algorithm would work well under good illumination. However, along with advantages, there are implications of good illumination too. Good illumination (a sunny day) and a decent sensor size, make it possible to capture images with a decreased exposure time. Thus, motion blur can be minimized and sharp images can be extracted where this decreases the noise in the disparity map.

Sharp images can capture all the details on the road, including the artefacts on the road which are noise for lane detection algorithms. Due to surface reflectivity (i.e. after rain) and the material of the asphalt (gravel, for example), large speckle noise can be introduced to the edge map. For the input image shown in 7(a), the edge map is estimated as shown in 7(b). The edge map is created with using the Sobel edge detector with a threshold of 80. It is seen that under the conditions mentioned above, large speckle noise is introduced. Using a higher edge threshold decreases this noise. However, with a high edge threshold under a shadow, lane edges also disappear from the edge map. For minimizing speckle noise, a well known median filter is a good option. The problem with a median filter is its fixed kernel size. If the kernel size is too small, noise elimination would be limited. On the other hand, if the kernel size is too large, lane features on the far field would be also eliminated. Thus, in this paper, a variable kernel size median filter is used. Consequently, used median filter has a large size in the near field and its size decreases directly proportional to the estimated road disparity value for each image row. Fig. 7(c) shows the output of the edge detection after filtering the input image with a small kernel size median filter. Fig. 7(d) shows the output of the edge detection after filtering the input image with a large kernel size median filter and Fig. 7(e) shows the output of the edge detection after filtering the input image with a variable kernel size median filter.

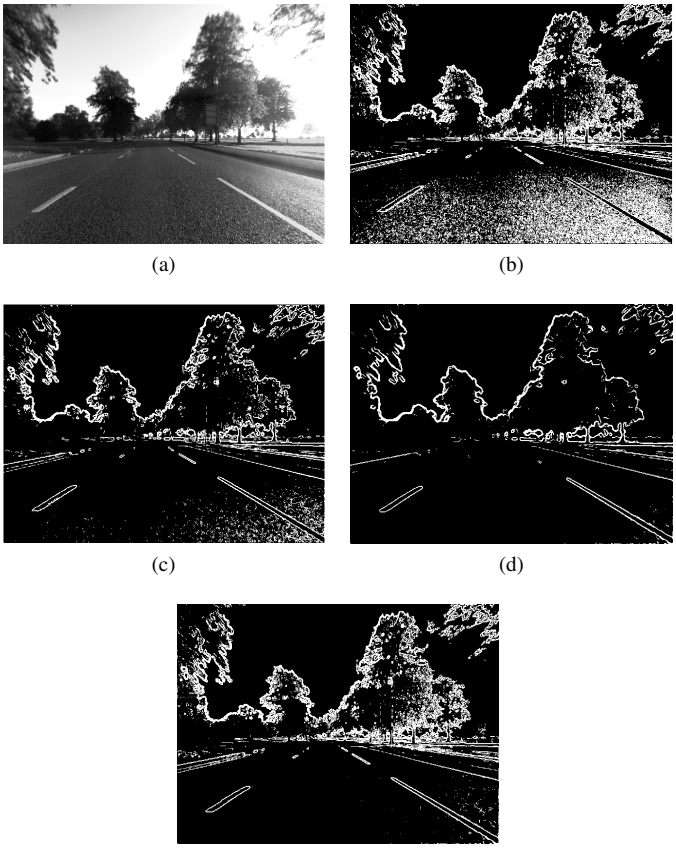
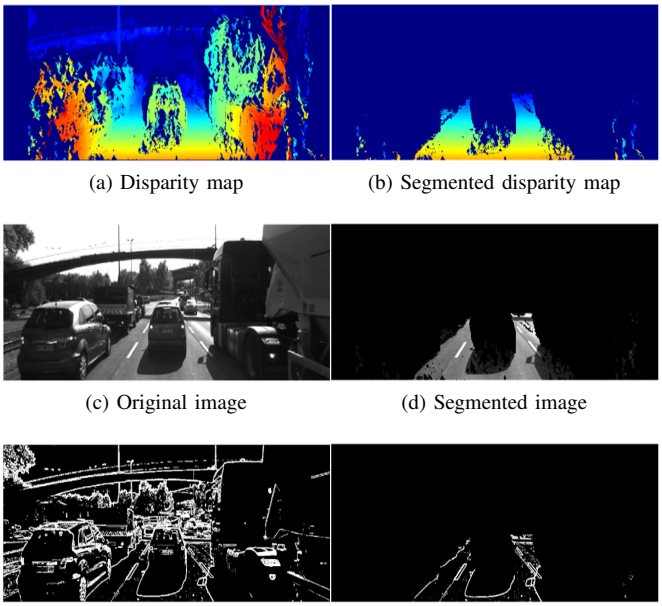


Fig. 7: Speckle noise filtering. (a) Input road image, (b) Edge map, (c) Edge map after filtering with small kernel size median filter, (d) Edge map after filtering with large kernel size median filter, (c) Edge map after filtering with variable kernel size median filter.

F. Road Segmentation

The feature map used for the lane detection is the edge map. For lane detection purposes, edges can be classified into a few categories: edges on the sky, the obstacles, the lane markings, the road markings and noise on the road such as caused by cracks and shadows. Since the vertical profile of the road is already calculated, this information can be used to eliminate further noise on the edge map and increase the SNR before calculating the horizontal profile of the road. Since some noise such as road cracks, shadows and road markings (apart from lane markings) appear on the road and



(a) Disparity map (b) Segmented disparity map

(c) Original image (d) Segmented image

(e) Edge map (f) Segmented edge map

Fig. 8: Feature map segmentation

they have the same disparity as the road, such noise cannot be eliminated by using the vertical profile of the road. However, most of the edge points caused by the obstacles and sky can be segmented. This is done by eliminating the edge points which have different disparity values (more than a few pixels) from the ones calculated by dynamic programming for their rows in the v-disparity map. Especially for the urban environment, this process can dramatically improve the SNR of the edge map, as seen in Fig. 8.

In Fig. 8(a), the calculated disparity map is shown. In Fig. 8(b), the calculated disparity is segmented for the road. In Fig. 8(c), the original image is shown. In Fig. 8(d), the original image is segmented by using the segmented disparity map. In Fig. 8(e), the original edge map is shown and, in Fig. 8(f), the edge map is segmented by using the segmented disparity map.

G. Horizon Line Calculation for Each Row of the Image

The horizon line (V_{py}) estimation is an important step for V_p detection. Some algorithms assume it is fixed and can be estimated by camera parameters [7] (ignoring the camera shakes) and some algorithms estimate a single horizon line [5]. For a flat road, a single horizon line can be estimated. However, for a non-flat road, the horizon line is continually changing according to the elevation of the road. Some algorithms [20], [14], propose a solution to the pitch angle estimation based on stereo vision which is essentially the same as estimating the horizon line for a flat road. Projection of the flat road to the v-disparity map is a straight line since the disparity of the road should decrease linearly and the road profile in this case can be estimated by using straight line detectors such as the Hough transform. However, for a non-flat road, the projection of the road is not a straight line. The approach taken in this paper is to estimate the horizon line by taking two points with different row numbers (their disparity values are already estimated using dynamic programming as in the section **??**) and, since they are close to each other, the change in elevation is small and that piece of the road can be assumed to be flat. By using these two points, a line equation can be calculated and the cross section of this lane and the column on v-disparity map which has 0 disparity value is the horizon line for that section of the road. After calculating the horizon line for this section, another 2 points are taken which are shifted one row above and the same process is applied to the estimated horizon line for this section. This process is iteratively calculated until the last estimated row of the road.

H. Forming Accumulator for V_{px}

V_p is composed of two values, V_{px} and V_{py}. In the previous sections, V_{py} is already estimated for each individual row of the image. The initial step of the proposed approach is to take a segment from the near field of the image and to form an accumulator for V_{px}. Then, all the edge points are voted to this accumulator according to each edge point's orientation and position. This initial step (forming the 1D V_{px} accumulator) is similar to the method used in our previous paper [7], In this work, instead of using a fixed horizon line, it is estimated by relying on the 3D information acquired by stereo vision.

The proposed approach then shifts the current band slightly up and creates another vanishing point accumulator. Computational efficiency is achieved by adding the edge point's votes which appear one row above the current band to the initially calculated accumulator and subtracting the edge point votes which appear on the bottom row of the current band from the initially calculated accumulator.

Furthermore, the far field of the road may contain a higher curvature. So, thinner bands are desirable for the upper bands of the image. Formation of thinner bands is achieved during the shifting process by subtracting more than one row from the bottom of the previous band, while adding only one row to the top of the previous band. As a result, the initial band to final band thickness ratio is adjustable.

By this approach, instead of creating an accumulator from scratch for every band, the algorithm updates the previous one by only processing and adding the top row (the vote positions of the bottom band are already calculated when adding them. Only subtraction is needed). Once an accumulator is updated for a band, the algorithm saves it to a 2D accumulator.

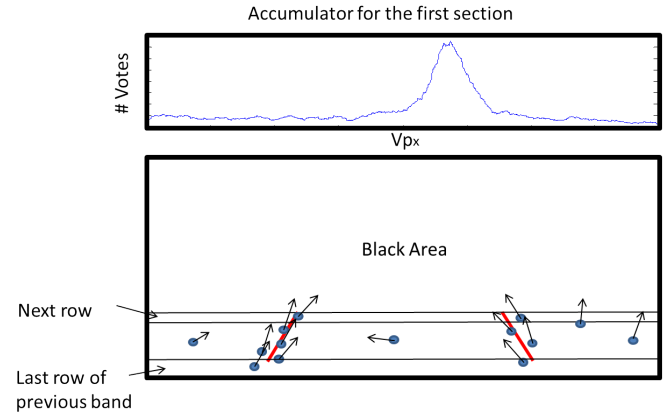
This process is demonstrated in Fig. 9. In Fig. 9(a), the input image is shown. In Fig. 9(b), the extracted feature image is shown. In Fig. 9(c), the process of creating 2D V_{px} accumulator is demonstrated. In Fig. 9(d), the extracted 2D V_{px} accumulator is shown (blue line in Fig. 9(d) is the estimated $V_{px}s$ and estimation method will be introduced in next section)

I. Estimating V_{px}

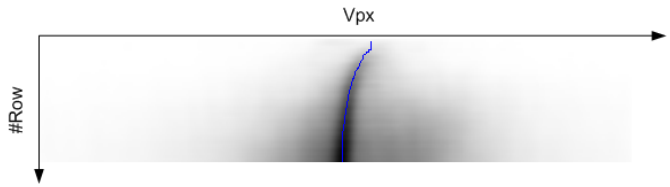
The vanishing point can be described as the cross section point between the tangent line of the lane and the horizon line. Thus, if the road model is a polynomial with an order of

(a) Original Image

(b) Feature map



(c) Shifting horizontal band for the V_{px} accumulator



(d) Resultant 2D V_{px} accumulator

Fig. 9: Creating the 2D V_{px} accumulator

n, then, the projection of this road model should also be the same order polynomial, as shown in equation 6,

$$Vpx(r) = u(r) - \frac{d}{dr}u(r) \cdot (r - Hz) \tag{6}$$

where $u(r)$ is the equation of the lane model. Thus, we can assume the vanishing point is changing gradually instead of having sudden jumps. In previous sections, the 2D V_{px} accumulator has already been constructed. Similar to the section II-D, this 2D accumulator can be optimized for a gradually changing output by dynamic programming to estimate the best path and acquiring a V_{px} for each row. For this purpose, equation 5 can be rewritten as equation 7.

$$E(Vpx) = m(Vpx_{r_{max}}) + \sum_{r=r_{max}-1}^{1} m(Vpx_r) + \lambda s(Vpx_r, Vpx_{r+1}) \tag{7}$$

For the feature map in Fig. 9(b), the resultant 2D V_{px} accumulator is shown in Fig. 9(d). In Fig. 9(d), a 2D V_{px} versus row position accumulator is demonstrated. The more votes a cell gets the darker it seems and the blue line on the top of the figure is the optimization result for this accumulator.

By scanning images as described in the previous sections, series of V_p is estimated and each calculated vanishing point is estimated as for the row which is in the middle of the corresponding band. In this way, each V_p is estimated except the rows under the middle of the first band. However, these rows are in the near field where the lanes tend to be straight and they can be estimated as the same as the Vp of the first band.

III. Forming The Likelihood Function

In the previous sections, a series of V_p is estimated. V_p can give direction and curvature information of the lanes. However, lateral positions of the lanes are still remain unknown. There is only one unknown variable left to detect for the lane detection. A 1D accumulator has been formed with a width of 2*image width. This approach is similar to the paper presented by [4]. For each possible value of intersection point (Constructed candidate lane and bottom row of the image), a likelihood value is calculated by allowing edge points underneath the constructed lane vote for individual starting points. Each edge point (e) votes according to the following equation;

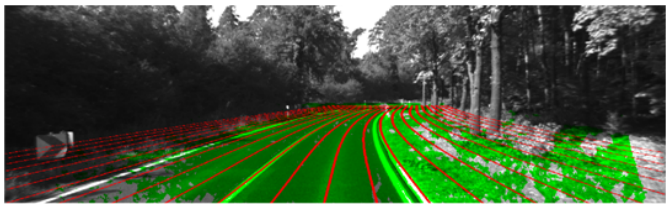
$$V(e) = \nabla(e) \cdot \cos(\theta_e - \theta_{Vp}) \tag{8}$$

where, for each individual edge point, V is the vote, $grad$ is the gradient, θ_e is the angle of the edge point and θ_{Vp} is the angle between the edge point and the vanishing point (in the future, inclusion of the connectivity is also planned). For the image in Fig. 10(a), the green area is the segmented road area and only the edge points in this area are used for voting. The red lines on Fig. 10(a) are created road patterns for some example starting points. In Fig. 10(b), the estimated 1D accumulator is illustrated for the image in Fig. 10(a).

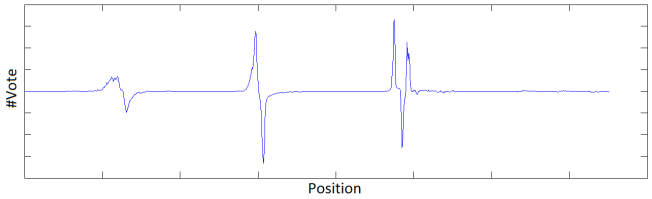
IV. Peak Pair Selection

Due to the dark-light-dark transition of the lane markings, a lane marking is projected into a 1D likelihood accumulator as a plus-minus peak pair. This property can be seen in Fig. 10(b).

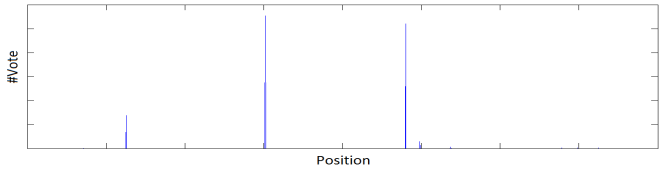
To detect these peak pairs, the algorithm initially finds plus and minus peak points and, secondly, for each plus peak point, finds minus peak points within a range and creates a Dirac function at the middle of the peak-pair in another accumulator. This process can be seen in Fig. 10. For the initial accumulator in Fig. 10(b), a new accumulator is created for peak pairs, as is seen in Fig. 10(c). Then, lateral offsets of the lanes are



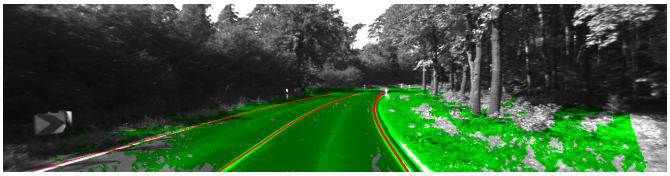
(a) For an example input image, the ground plane is illustrated with the green area (excluding obstacles, sky and occluded areas in disparity map) and some of the candidate lane markings are illustrated with red lines.



(b) Projection of candidate lane markings to the 1D accumulator (size of this signal is two times larger than the image width).



(c) 1D signal after plus-minus peak pair selection



(d) Lane detection result

Fig. 10: Creating 1D accumulators

selected from this new accumulator. Starting from the highest peak, the algorithm eliminates all other peaks within the range ($\pm 1m$) and, then, detects the next highest peak from the signal and eliminates all other peaks within the range ($\pm 1m$). The algorithm, iterates until the detected peak is lower then the selected threshold. The detected lanes for Fig. 10(b) are illustrated in Fig. 10(d).

V. 1D Signal Noise Reduction

The proposed algorithm can detect multiple lanes. However, compared to the lane in which the vehicle is travelling, the lanes further from the vehicle are harder to detect, especially when these lanes are dashed. To detect multiple lanes consistently, a noise reduction step is applied. In this paper, via the estimated Vps, the lane detection is reduced to a 1D problem. Let the noise-free signal be x(n) and the estimated signal be:

$$S_i(n) = x(n) + w(n) \tag{9}$$

where $w(n)$ is Gaussian noise. For instance, $w(n)$ can be artefacts on the road which are directed to the Vp but, however, do not consistently appear on consecutive frames. If the road thicknesses are the same for two consecutive frames, the noise-free signal for frame $i+1$ would be $x(n-L)$ where L is the change in lateral offset. Thus, the estimated signal for frame $i+1$ can be defined as in equation 10;

$$S_{i-1}(n) = x(n-L) + w(n) \tag{10}$$

Thus, the change in lateral offset (L) between two consecutive frames can be estimated by applying cross-correlation between these two signals via equation 11. In this way, before detecting the lanes, the lateral offset change can be estimated.

$$r_{i,i-1}(k) = \sum_{n=0}^{N} S_i(n) \cdot S_{i-1}(n+k) \tag{11}$$

The probability density function (PDF) of the change in lateral offset (L) is then estimated by normalizing $r_{i,i+1}(k)$ for the range of k.

$$PDF_{i,i-1}(k) = r_{i,i-1}(k)/T \tag{12}$$

where

$$T_{i,i-1}(k) = \sum_{k=-20}^{20} r_{i,i-1}(k) \tag{13}$$

Finally, signal alignment can be achieved by convolving PDF of L with the next signal. For random noise reduction (improvement in SNR), a few signals for a few consecutive frames are iteratively aligned and added together as in the equation below.

$$ST_i(n) = S_i(n) + S_{i-1}(n) * PDF_{i,i-1} \tag{14}$$

VI. Experimental Results

To quantify the performance of the algorithm, we tested both the accuracy of V_p and the lane detection ratio. For estimating the error in V_p, first the ground truth V_p locations should be known. For estimating the ground truth V_p, selecting a V_p location manually from the image would be inaccurate. In this paper, for a more accurate estimate of V_p, the user selected 4 points manually, 2 from one lane and 2 from another lane. Then, the script first calculates 2 line equations and, then, using the estimated line parameters, calculates the cross section point (V_p) of these 2 lines. In this way, V_p can be estimated accurately. However, this process needs manual work and thus the ground truth V_p locations are only estimated for the near field.

Once the ground truth positions of V_ps are estimated, the mean error of V_p is calculated for the sample sequences. Then, the mean error of V_p for the algorithm described in [7] is estimated for comparison. The problem with this algorithm is it needs V_{py} as input. Thus, the ground truth V_{py} is given as the input and the mean error in V_{px} is estimated. The detailed results are shown in Table I.

To quantify the robustness of the system, the detection ratio of the algorithm is also estimated on sample sequences from KITTY datasets [21]. Sample detection results are illustrated

TABLE I: Mean error in vanishing points

	Proposed algorithm		Y.Wang et. al. [7]	
	Err. Vpy	Err. Vpx	Err. Vpy	Err. Vpx
Sequence 1 (429 frame)	4.4825	4.4916	Ground truth input	10.4558
Sequence 2 (295 frame)	5.0319	4.8418	Ground truth input	9.5621
Sequence 3 (187 frame)	3.8749	3.3689	Ground truth input	14.4960

TABLE II: Detection results

Sequence	Total Lane markings	Correct detection	Incorrect detection	Mis-Detectin
Sequence 1	860	860	0	0
Sequence 2	594	594	0	0
Sequence 3	276	276	0	0
Sequence 4	156	147	0	9
Sequence 5	678	661	0	17
total	2564	2538	0	26

in Fig. 11 and the detailed results (sample video sequences will be available at http://ieeexplore.ieee.org) are shown in Table II (only the closest lanes are taken into account in this table).

VII. Future Work

In this paper, the algorithm searches for dark-light-dark patterns for the lanes which is a more robust way (compared to searching for a single boundary) to detect lanes and currently it can only detect painted lanes. In Fig. 10.a, it can be seen that there are both lane marking and a road boundary on the right hand side of the road. This is projected to the 1D likelihood accumulator, as in Fig. 10.b, as peaks. In that position, there is a plus peak just after the plus-minus peak pair, where the plus-minus peak pair is due to the lane marking and the plus peak just after this peak pair is due to the road boundary. If there were not any painted lane markings, there would be only a positive peak. To detect this road boundary, it is possible to modify the peak pair selection in an ad-hoc manner. Such as, if no lane is detected on either side of the road, the algorithm should also search for the single high peaks. Thus, the algorithm would be able to detect painted lanes robustly by using the dark-light-dark pattern of the lanes and, at the same time, it would be able to detect road boundaries when, painted lane markings are not available.

Currently, the algorithm detects lanes accurately and robustly. However, further improvement is possible by applying tracking. For this purpose, it is also planned to fit estimated V_{px} values and V_{py} values of a frame into spline models and track estimated control points. Thus, tracking lanes would be possible by tracking few parameters.

VIII. Conclusion

In this paper, a novel lane detection and tracking algorithm is presented. The main novel elements of this paper include dense vanishing point estimation, the use of estimated vanishing points to detect lanes, estimating the change in lateral offset of the car in a global way and utilizing this estimation for SNR improvement. The V_p contains the global information of the road image. Hence, V_p based lane detection algorithms are quite insensitive to interference and they demonstrate robust results. Traditionally, V_p based lane detection algorithms deal with a single V_p in the whole image under the assumptions of a straight lane and a flat road. Although, previously using multiple V_{px} is also proposed by some authors, these algorithms still have a flat road assumption (they use a single V_{py}) and suggested V_{px} estimation techniques are not global. Unlike the previously proposed V_p based lane detection algorithms, the algorithm described in this paper proposes a global approach for dense vanishing point estimation and can detect multiple lanes robustly and accurately with both horizontal and vertical curvature. Experimental results show that, proposed algorithm works robustly and accurately even in dense traffic. Furthermore, due to the flexibility of the described system, the user can simply plug a stereo camera rig (experimental set-up) on to a vehicle without concern about any of the external camera parameters (i.e. camera height or pitch, yaw and roll angle).

References

[1] *Reported road casualties Great Britain 2009*. The Stationery Office Department for Transport, 2010.

[2] J. C. McCall and M. M. Trivedi, "Video-based lane estimation and tracking for driver assistance: survey, system, and evaluation," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 7, no. 1, pp. 20–37, 2006.

[3] A. B. Hillel, R. Lerner, D. Levi, and G. Raz, "Recent progress in road and lane detection: a survey," *Machine Vision and Applications*, pp. 1–19, 2012.

[4] D. Schreiber, B. Alefs, and M. Clabian, "Single camera lane detection and tracking," in *Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE*. IEEE, 2005, pp. 302–307.

[5] D. Hanwell and M. Mirmehdi, "Detection of lane departure on high-speed roads." in *ICPRAM (2)*, 2012, pp. 529–536.

[6] B. Fardi and G. Wanielik, "Hough transformation based approach for road border detection in infrared images," in *Intelligent Vehicles Symposium, 2004 IEEE*. IEEE, 2004, pp. 549–554.

[7] Y. Wang, N. Dahnoun, and A. Achim, "A novel system for robust lane detection and tracking," *Signal Processing*, vol. 92, no. 2, pp. 319–334, 2012.

[8] Y. Wang, E. K. Teoh, and D. Shen, "Lane detection and tracking using b-snake," *Image and Vision computing*, vol. 22, no. 4, pp. 269–280, 2004.

[9] R. Labayrade and D. Aubert, "A single framework for vehicle roll, pitch, yaw estimation and obstacles detection by stereovision," in *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*. IEEE, 2003, pp. 31–36.

[10] N. Lazaros, G. C. Sirakoulis, and A. Gasteratos, "Review of stereo vision algorithms: from software to hardware," *International Journal of Optomechatronics*, vol. 2, no. 4, pp. 435–462, 2008.

[11] N. Einecke and J. Eggert, "A two-stage correlation method for stereoscopic depth estimation," in *DICTA*, 2010, pp. 227–234.

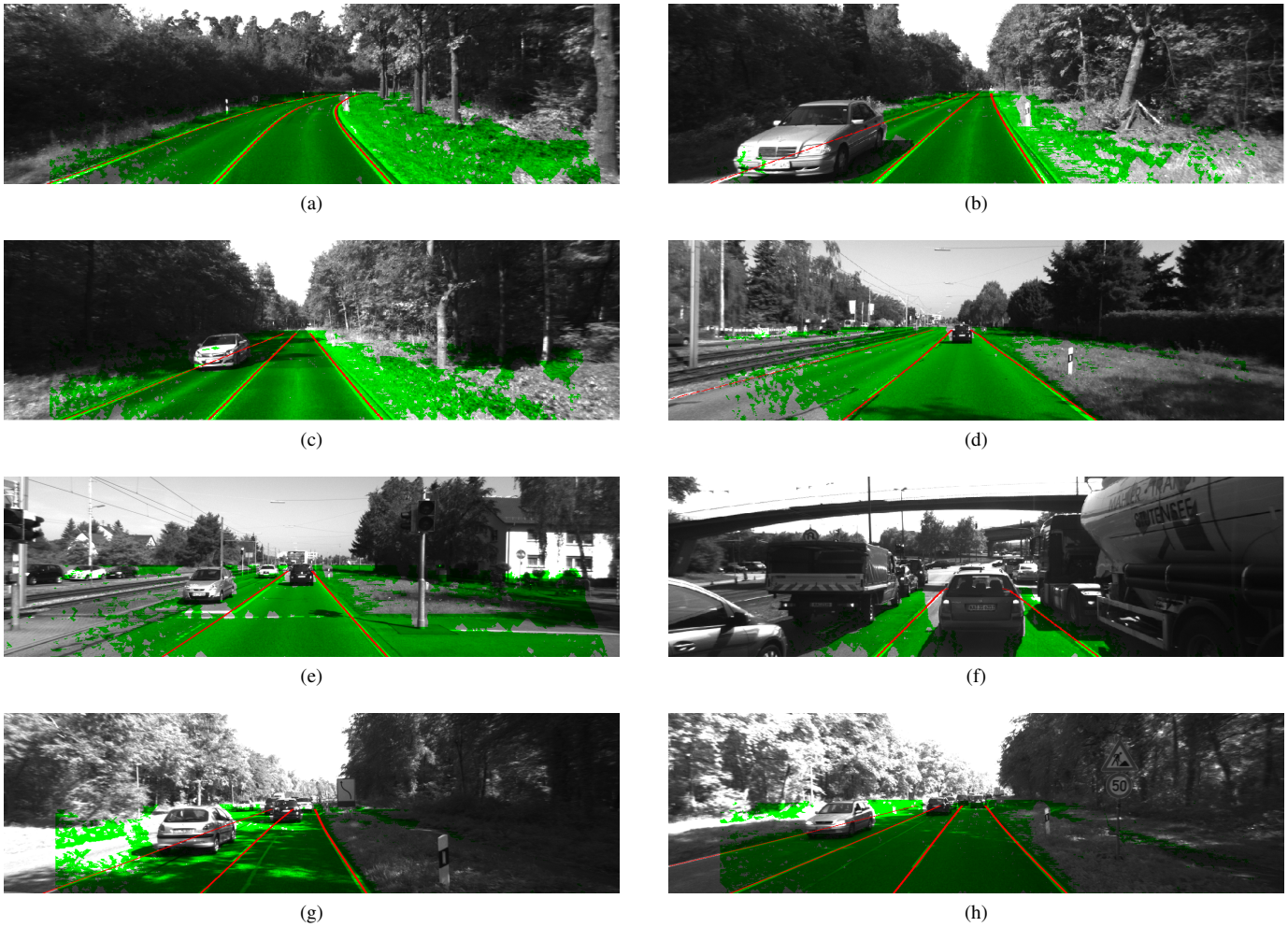
[12] A. Geiger, M. Roser, and R. Urtasun, "Efficient large-scale stereo matching," in *ACCV*, 2010.

[13] Z. Zhang, X. Ai, and N. Dahnoun, "Efficient disparity calculation based on stereo vision with ground obstacle assumption," in *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*. IEEE, 2013, pp. 1–5.

[14] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through v-disparity representation," in *Intelligent Vehicle Symposium, 2002. IEEE*, vol. 2. IEEE, 2002, pp. 646–651.

[15] Y. Gao, X. Ai, Y. Wang, J. Rarity, and N. Dahnoun, "U-v disparity based obstacle detection with 3d camera and steerable filter," in *Intelligent Vehicles Symposium (IV), 2011 IEEE*. IEEE, 2011, pp. 957–962.

[16] S. Nedevschi, R. Schmidt, T. Graf, R. Danescu, D. Frentiu, T. Marita, F. Oniga, and C. Pocol, "3d lane detection system based on stereovision," in *Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference on*. IEEE, 2004, pp. 161–166.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 11: Experimental Results

[17] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.

[18] P. F. Felzenszwalb and D. P. Huttenlocher, "Pictorial structures for object recognition," *International Journal of Computer Vision*, vol. 61, no. 1, pp. 55–79, 2005.

[19] O. Veksler, "Stereo correspondence by dynamic programming on a tree," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2. IEEE, 2005, pp. 384–390.

[20] A. D. Sappa, D. Gerónimo, F. Dornaika, and A. López, "Real time vehicle pose using on-board stereo vision system," in *Image Analysis and Recognition*. Springer, 2006, pp. 205–216.

[21] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.

Umar Ozgunalp received the BSc. degree in electrical and electronics engineering from the Eastern Mediterranean University, Famagusta, Cyprus, in 2007 and MSc. degree in electronic communications and computer engineering from the University of Nottingham, UK, in 2009. He is working toward the PhD. degree at the Visual Information Laboratory, electrical and electronics engineering, University of bristol, Bristol, UK. His research interests include computer vision and pattern recognition.

Xiao Ai is a postdoctoral researcher at Bristol University, having recently graduated from his PhD program which specialized in 3D imaging techniques. Xiao 's research involves the development of lidar technologies: for atmospheric sensing and automotive obstacle detection. Xiao also has extensive experience on optoelectronic systems, analogue and digital signal processing and machine vision.

N. Dahnoun Reader in the Electrical and Electronic Engineering at the University of Bristol. He obtained his PhD in Biomedical Engineering in 1990 and is the Programme Director for Computer Science and Electronics at the University of Bristol. Naim has over 25 years experience in real-time DSP and, in recognition of the important role played by Universities in educating engineers in new technologies such as real-time Digital Signal Processing (DSP).Naim 's main research interests include Real-Time Digital Signal Processing applied to Biomedical Engineering, Video Surveillance, Automotive and Optics.